

SpringerBriefs in Computer Science

Series Editors

Stan Zdonik, Brown University, Providence, RI, USA

Shashi Shekhar, University of Minnesota, Minneapolis, MN, USA

Xindong Wu, University of Vermont, Burlington, VT, USA

Lakhmi C. Jain, University of South Australia, Adelaide, SA, Australia

David Padua, University of Illinois Urbana-Champaign, Urbana, IL, USA

Xuemin Sherman Shen, University of Waterloo, Waterloo, ON, Canada

Borko Furht, Florida Atlantic University, Boca Raton, FL, USA

V. S. Subrahmanian, University of Maryland, College Park, MD, USA

Martial Hebert, Carnegie Mellon University, Pittsburgh, PA, USA

Katsushi Ikeuchi, University of Tokyo, Tokyo, Japan

Bruno Siciliano, Università di Napoli Federico II, Napoli, Italy

Sushil Jajodia, George Mason University, Fairfax, VA, USA

Newton Lee, Institute for Education, Research and Scholarships, Los Angeles, CA, USA

SpringerBriefs present concise summaries of cutting-edge research and practical applications across a wide spectrum of fields. Featuring compact volumes of 50 to 125 pages, the series covers a range of content from professional to academic.

Typical topics might include:

- A timely report of state-of-the art analytical techniques
- A bridge between new research results, as published in journal articles, and a contextual literature review
- A snapshot of a hot or emerging topic
- An in-depth case study or clinical example
- A presentation of core concepts that students must understand in order to make independent contributions

Briefs allow authors to present their ideas and readers to absorb them with minimal time investment. Briefs will be published as part of Springer's eBook collection, with millions of users worldwide. In addition, Briefs will be available for individual print and electronic purchase. Briefs are characterized by fast, global electronic dissemination, standard publishing contracts, easy-to-use manuscript preparation and formatting guidelines, and expedited production schedules. We aim for publication 8–12 weeks after acceptance. Both solicited and unsolicited manuscripts are considered for publication in this series.

****Indexing:** This series is indexed in Scopus, Ei-Compendex, and zbMATH ******

More information about this series at <https://link.springer.com/bookseries/10028>

Youyang Qu · Longxiang Gao · Shui Yu ·
Yong Xiang

Privacy Preservation in IoT: Machine Learning Approaches

A Comprehensive Survey and Use Cases

Youyang Qu 
Data61
Australia Commonwealth Scientific
and Industrial Research Organization
Melbourne, VIC, Australia

Shui Yu 
School of Computer Science
University of Technology Sydney
Ultimo, NSW, Australia

Longxiang Gao 
Shandong Computer Science Center
Qilu University of Technology
Shandong, China

Yong Xiang 
School of Information Technology
Deakin University
Burwood, VIC, Australia

ISSN 2191-5768 ISSN 2191-5776 (electronic)
SpringerBriefs in Computer Science
ISBN 978-981-19-1796-7 ISBN 978-981-19-1797-4 (eBook)
<https://doi.org/10.1007/978-981-19-1797-4>

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

Internet of Things (IoT), as a booming computing architecture, is experiencing rapid development with a speed beyond imagination. Nowadays, IoT devices are so pervasive that they have become key components of human daily life, such as sensors, intelligent cameras, smart wearable devices, and a lot more. By reshaping the existing network architecture, IoT has provided significant convenience and improvement of quality of life.

Since IoT devices are deployed ubiquitously, an increasing volume of data is collected and transmitted over IoTs. The statistic shows total data volume of connected IoT devices worldwide is forecast to reach 79.4 zettabytes (ZBs) by 2025. However, the data privacy issues become even severe because sensitive information of collected data is not properly managed, especially health data, location data, identity-related data, etc. Moreover, data from multiple sources pose further challenges since the interconnections among the data may reveal more sensitive information. Furthermore, the advancement of data pattern extraction and data analysis techniques put privacy under more serious threats. Thus, privacy preservation has become a crucial issue that needs to be well considered in this age of IoT.

Machine learning has proved its superior performance in data manipulation field. In addition to perform predictive analysis or optimization-oriented services, machine learning algorithms are adopted in privacy-preserving data sharing and publishing scenarios. It attracts extensive interest from both academia and industry. Among all existing solutions, reinforcement learning, federated learning, and generative adversarial networks (GAN) are the most popular and practical ones. Extensive research has been conducted to leverage or modify them for privacy protection considering diverse conditions. Therefore, they are also the main focus of this monograph, through which the rationale of machine-learning-driven privacy protection solutions are present.

In this monograph, we are going to comprehensively and systematically introduce machine-learning-driven privacy preservation in Internet of Things (IoT). In this big data era, an increasingly massive volume of data is generated and transmitted in IoTs, which poses great threats to privacy protection. Motivated by this, an emerging research topic, machine-learning-driven privacy preservation, is fast

booming to address various and diverse demands of IoTs. However, there is no existing literature discussion on this topic in a systematical manner. The authors in this monograph aim to sort out the clear logic of the development of machine-learning-driven privacy preservation in IoTs, the advantages, and disadvantages of it, as well as the future directions in this under-explored domain. The issues of existing privacy protection methods (differential privacy, clustering, anonymity, etc.) for IoTs, such as low data utility, high communication overload, and unbalanced trade-off, are identified to the necessity of machine-learning-driven privacy preservation. Besides, the leading and emerging attacks pose further threats to privacy protection in this scenario. To mitigate the negative impact, machine-learning-driven privacy preservation methods for IoTs are discussed in detail on both the advantages and flaws, which is followed by potentially promising research directions.

The prominent and exclusive features of this book are as follows:

- Reviews exhaustive the key recent research into privacy-preserving techniques in IoTs.
- Enriches understanding of emerging machine learning enhanced privacy-preserving techniques in IoTs.
- Covers several real-world applications scenarios.
- Maximize reader insights into how machine learning can further enhance privacy protection in IoTs.

This monograph aspires to keep readers, including scientists and researchers, academic libraries, practitioners and professionals, lecturers and tutors, postgraduates, and undergraduates, updated with the latest algorithms, methodologies, concepts, and analytic methods for establishing future models and applications of machine-learning-driven privacy protection in IoTs. It not only allows the readers to familiarize with the theoretical contents but also enables them to make best use of the theories and develop new algorithms that could be put into practice.

The book contains roughly three main modules. In the first module, the book presents how to achieve decentralized privacy using blockchain-enabled federated learning. In the second module, the personalized privacy protection model using GAN-driven differential privacy is given. In the third module, the book shows the hybrid privacy protection using reinforcement learning. Based on the above knowledge, the book presents the identified open issues and several potentially promising future directions of personalized privacy protection, followed by a summary and outlook on the promising field. In particular, each of the chapter is self-contained for the readers' convenience. Suggestions for improvement will be gratefully received.

Melbourne, Australia
Shandong, China
Ultimo, Australia
Burwood, Australia

Youyang Qu
Longxiang Gao
Shui Yu
Yong Xiang

Acknowledgments

We sincerely appreciate numerous colleagues and postgraduate students at Deakin University, Melbourne and University of Technology Sydney, Sydney, who contribute a lot from various perspectives such that we are inspired to write this monograph. We would like to acknowledge the support from the research grant we received, namely, ARC Discovery Project under the file number of 200101374. In this book, some interesting research results demonstrated are extracted from our research publications that indeed (partially) supported through the above research grants. We are also grateful to the editors of Springer, especially Dr. Nick Zhu, for his continuous professional support and guidance. Finally, we would like to express our thanks to the family of each of us for their persistent and selfless supports. Without their encouragement, the book may regrettably become some fragmented research discussions.

Melbourne, Australia
Shandong, China
Sydney, Australia
Melbourne, Australia
December 2021

Youyang Qu
Longxiang Gao
Shui Yu
Yong Xiang

Contents

1	Introduction	1
1.1	IoT Privacy Research Landscape	1
1.2	Machine Learning Driven Privacy Preservation Overview	2
1.3	Contribution of This Book	3
1.4	Book Overview	4
2	Current Methods of Privacy Protection in IoTs	7
2.1	Briefing of Privacy Preservation Study in IoTs	7
2.2	Cryptography-Based Methods in IoTs	9
2.3	Anonymity-Based and Clustering-Based Methods	11
2.4	Differential Privacy Based Methods	13
2.5	Machine Learning and AI Methods	14
2.5.1	Federated Learning	15
2.5.2	Generative Adversarial Network	16
	References	16
3	Decentralized Privacy Protection of IoTs Using Blockchain-Enabled Federated Learning	19
3.1	Overview	19
3.2	Related Work	21
3.3	Architecture of Blockchain-Enabled Federated Learning	23
3.3.1	Federated Learning in FL-Block	23
3.3.2	Blockchain in FL-Block	24
3.4	Decentralized Privacy Mechanism Based on FL-Block	27
3.4.1	Blocks Establishment	27
3.4.2	Blockchain Protocols Design	29
3.4.3	Discussion on Decentralized Privacy Protection Using Blockchain	29
3.5	System Analysis	30
3.5.1	Poisoning Attacks and Defence	30
3.5.2	Single-Epoch FL-Block Latency Model	31
3.5.3	Optimal Generation Rate of Blocks	34

3.6	Performance Evaluation	35
3.6.1	Simulation Environment Description	35
3.6.2	Global Models and Corresponding Updates	37
3.6.3	Evaluation on Convergence and Efficiency	38
3.6.4	Evaluation on Blockchain	41
3.6.5	Evaluation on Poisoning Attack Resistance	42
3.7	Summary and Future Work	45
	References	45
4	Personalized Privacy Protection of IoTs Using GAN-Enhanced Differential Privacy	49
4.1	Overview	50
4.2	Related Work	51
4.3	Generative Adversarial Nets Driven Personalized Differential Privacy	53
4.3.1	Extended Social Networks Graph Structure	53
4.3.2	GAN with a Differential Privacy Identifier	54
4.3.3	Mapping Function	57
4.3.4	Optimized Trade-Off Between Personalized Privacy Protection and Optimized Data Utility	61
4.4	Attack Model and Mechanism Analysis	62
4.4.1	Collusion Attack	62
4.4.2	Attack Mechanism Analysis	63
4.5	System Analysis	64
4.6	Evaluation and Performance	65
4.6.1	Trajectory Generation Performance	67
4.6.2	Personalized Privacy Protection	67
4.6.3	Data Utility	69
4.6.4	Efficiency and Convergence	69
4.6.5	Further Discussion	71
4.7	Summary and Future Work	73
	References	74
5	Hybrid Privacy Protection of IoT Using Reinforcement Learning	77
5.1	Overview	78
5.2	Related Work	80
5.3	Hybrid Privacy Problem Formulation	80
5.3.1	Game-Based Markov Decision Process	80
5.3.2	Problem Formulation	81
5.4	System Modelling	82
5.4.1	Actions of the Adversary and User	82
5.4.2	System States and Transitions	83
5.4.3	Nash Equilibrium Under Game-Based MDP	84

- 5.5 System Analysis 86
 - 5.5.1 Measurement of Overall Data Utility 86
 - 5.5.2 Measurement of Privacy Loss 86
- 5.6 Markov Decision Process and Reinforcement Learning 88
 - 5.6.1 Quick-Convergent Reinforcement Learning Algorithm 88
 - 5.6.2 Best Strategy Generation with Limited Power 89
 - 5.6.3 Best Strategy Generation with Unlimited Power 90
- 5.7 Performance Evaluation 91
 - 5.7.1 Experiments Foundations 92
 - 5.7.2 Data Utility Evaluations 92
 - 5.7.3 Privacy Loss Evaluations 96
 - 5.7.4 Convergence Speed 102
- 5.8 Summary and Future Work 104
- References 106
- 6 Future Research Directions 111**
 - 6.1 Trade-Off Optimization in IoTs 111
 - 6.2 Privacy Preservation in Digital Twined IoTs 112
 - 6.3 Personalized Consensus and Incentive Mechanisms
for Blockchain-Enabled Federated Learning in IoTs 113
 - 6.4 Privacy-Preserving Federated Learning in IoTs 113
 - 6.5 Federated Generative Adversarial Network in IoTs 114
- 7 Summary and Outlook 117**